

BERNHIST: Eine raum-zeitliche Datenbank für den Schweizer Kanton Bern im Internet

Imfeld, Klaus; Pfister, Christian; Häberli, Peter

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Imfeld, K., Pfister, C., & Häberli, P. (1995). BERNHIST: Eine raum-zeitliche Datenbank für den Schweizer Kanton Bern im Internet. *Historical Social Research*, 20(3), 102-133. <https://doi.org/10.12759/hsr.20.1995.3.102-133>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/4.0>

DATA & ARCHIVES

BERNHIST

Eine raum-zeitliche Datenbank für den Schweizer Kanton Bern im Internet

*Klaus Imfeld, Christian Pfister, Peter Häberli**

Abstract: A historic-geographical relational data-base named Bernhist was created for the Swiss canton of Bern (population 1 million) over the last decade. It includes statistical sources from the period 1700 to 1990 that are stored in three-dimensional tables including space, time and topic. The tool allows examining the spatial distribution of population and resources for different points in time. The geographical structure (27 districts, 410 communities) and the data dictionary are tailored according to those of the Swiss Statistical Office in order to keep the historical evidence compatible to that of the present. The data-base comprises two million entities. The focus is on population (including a complete survey of births and deaths), agriculture, occupation and social indicators. Bernhist can be accessed directly on the Internet (address 'ubleclu.unibe.ch') from every PC with standard terminal software capable to emulate the DEC VT200 standard. The account name is 'uni-info', no password. As the data-base has to meet the needs of a mixed community of users including nonprofessional historians (people from the media and from the administration, high school students etc.) the queries only involve a few keys and some easy commands in German. The results have the form of a three dimensional table that can be arranged in six different ways. Output may be dumped to a

* Address all communications to Klaus Imfeld, Christian Pfister, Peter Häberli, Forschungsstelle für Regional- und Umweltgeschichte, Historisches Institut der Universität Bern, Unitobler, CH-3000 Bern 9, e-mail: bhist@hist.unibe.ch.

flat ASCII file (licensed PC's only) and thence may be easily imported into other (Statistical, cartographic) Software for analysis. Bernhist provides the means, through systematic querying, to discover associations between data items that might otherwise pass unnoticed or be impossible to find. An introductory monograph (Pfister 1995) is being published.

1. Einführung

Die gängige Forderung, Probleme seien fächerübergreifend anzugehen und Zusammenhänge zwischen verschiedenartigen Phänomenen und Prozessen seien aufzuzeigen, kann in der Praxis der (landes-) geschichtlichen Forschung längst nicht immer eingelöst werden: Oft zwingen arbeitsökonomische Anforderungen zu einer Beschränkung auf Teilaspekte; häufiger noch lassen sich die erforderlichen Daten für den gewählten Untersuchungszeitraum nicht mit einem vernünftigen Zeitaufwand beschaffen, oder sie sind gar nicht erhoben worden. Leistungsfähige Datenbanken, die in dieser Situation Hilfestellung bieten könnten, stehen für die Geschichtswissenschaft noch nicht zur Verfügung.

In den letzten zehn Jahren ist für den Kanton Bern eine relationale Datenbank namens BERNHIST¹ aufgebaut worden, aus der eine neue, vernetzte Landesgeschichte herauswachsen könnte. BERNHIST enthält einen so umfangreichen und vielfältigen Bestand an historisch-statistischen Quellen, daß (teil-) fächerinterne wie -übergreifende Fragestellungen ohne großen Suchaufwand angegangen werden können. Derzeit umfaßt das Gefäß rund zwei Millionen einzelne Entitäten (»Zahlen«) aus den Bereichen Bevölkerung, Landwirtschaft, Gewerbe/Industrie, Umwelt, Sozialstruktur und Politik auf drei räumlichen Ebenen (Gemeinden, Bezirke, Gesamtkanton) für den Zeitraum von 1700 bis zur Gegenwart.

Demnächst wird eine öffentliche Version von BERNHIST für die ortsungebundene Anwendung in Lehre, Forschung, Dienstleistung und Öffentlichkeitsarbeit freigegeben. Die Datenbank kann - auf der Basis von MS-DOS - von überall her abgerufen werden, wo ein Anschluß ans Telekommunikationsnetz Internet besteht².

Bei der Einrichtung der heutigen Version der Datenbank war die Frage nach den mutmaßlichen Kunden und ihren möglichen Wünschen an Bedienungs-komfort und Information maßgebend. Dabei wurde angenommen, daß für Angehörige der Universität vor allem die Vielseitigkeit des Instruments im Vordergrund steht, während die in Schule, Verwaltung und im Medienbereich Tä-

¹ Pfister, C, Schüle, H., 1988, S. 229-267.

² Internet-Adresse: »ubecclu.unibe.ch« oder »130.92.4.11«, Username: Uniinfo, dann Menüauswahl.

tigen mehr Gewicht auf die eine benutzerfreundliche Oberfläche und die Möglichkeit zur raschen graphischen Aufbereitung legen. Für diese erweiterte Zielgruppe ist die Datenbank hauptsächlich konzipiert worden. Die Kosten für den Aufbau eines solchen Instruments rechtfertigen sich nur durch einen entsprechend großen Kundenkreis außerhalb der Universität. Daß es auch der universitären Forschung den Weg ebnet ist bloß ein zusätzlicher Grund.

Worin besteht die Besonderheit eines historisch-geographischen Informationssystems (HGIS) wie BERNHIST im Vergleich mit anderen GIS-Systemen? Die Datenbanken STATINF und GEOSTAT des Bundesamtes für Statistik schließen alle Bereiche der heutigen Statistik auf allen räumlichen Ebenen - Gesamtschweiz, Kantone, Bezirke, Gemeinden, bis hinunter zum Hektarraster - aber nur für die letzten Jahrzehnte. Die vor der Publikation stehende »Historische Statistik der Schweiz«¹ umfaßt die Periode des Bundesstaates und deckt in diesem Zeitraum die wichtigsten statistischen Bereiche ab; räumlich reicht sie jedoch nur bis auf die Ebene der Kantone. BERNHIST beruht räumlich wie STATINF auf der Basis der Einwohnergemeinden und greift zeitlich weiter in die Vergangenheit zurück als die »Historische Statistik«. Thematisch ist das Angebot von BERNHIST dagegen schmäler als jenes der anderen beiden Instrumente; die knappen Mittel zwangen zur Beschränkung auf Wesentliches. Der gesamte Datenbestand von BERNHIST ist in vier Gefäßen abgelegt (vgl. *Abb. 1*), die sich teilweise überlappen.

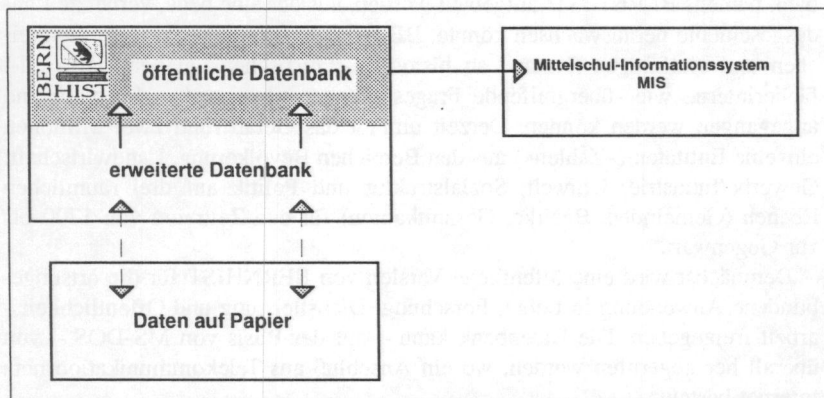


ABB. 1: *Der Datenbestand des Gesamtsystems BERNHIST, verteilt auf vier Gefäße. In diesem Text wird vor allem vom öffentlichen Teil die Rede sein.*

Die Datenbank selber besteht aus zwei Teilen. Ein über Internet zugänglicher öffentlicher Teil enthält jenes Material, von dem angenommen wird, daß es

¹ Ritzmann-Blickenstorfer, H. 1994.

vom Publikum in erster Linie nachgefragt wird. Manche dieser Dateien sind schon seit einiger Zeit ins Mittelschul-Informationssystem MIS* integriert, wo sie mit Modem abgerufen werden (BERNHIST auf dem MIS). Da alle bemschen Schulen in absehbarer Zeit mit einem Modem ausgerüstet werden sollen, werden diese Bestände ohne großen Aufwand für die Unterrichtsgestaltung fruchtbar gemacht werden können.

Neben Daten von öffentlichem Interesse sind im Laufe der Zeit Bestände erhoben und ausgewertet worden, die nur einen kleinen Kreis ansprechen. Der Einbezug dieses »erweiterten Teils« in den »öffentlichen Teil« hätte den Themenkatalog¹ übermäßig aufgebläht. Die Daten des »erweiterten Teils« können an Geräten der Forschungsstelle für Regional- und Umweltgeschichte abgefragt werden. Ein weiterer Teil der Bestände ist noch in Form der ursprünglichen SPSSX-Files² gespeichert und nicht in die Datenbank integriert. Er könnte in maschinenlesbarer Form nur mit einem erheblichen Programmieraufwand konsultiert werden. Dieser Teil ist vorläufig in Ordnern abgelegt (BERNHIST auf Papier)³. Die nicht öffentlich zugänglichen Bestände können nur nach Voranmeldung konsultiert werden, und der dazu benötigte Personalaufwand muß entschädigt werden.

Im folgenden wird die öffentliche Version der Datenbank BERNHIST vorgestellt. Inhaltliche Probleme werden dabei nicht angesprochen. Sie sind Gegenstand einer besonderen Publikation⁴. Die Entwicklung eines Datenbankmodells und dessen anschließende Implementierung auf einer realen Maschine sind informationstechnische Komplexe, die dem Laien nicht unbedingt verständlich sind. Es wurde versucht, diese Komplexität zu reduzieren und das Konzept von BERNHIST auch denjenigen zugänglich zu machen, die über keine Informatikkenntnisse verfügen. Eine umfassende Beschreibung des gesamten Systems ist in Vorbereitung⁵. Dieser Artikel soll das Konzept, den Aufbau und die Funktionsweise des Forschungsinstruments BERNHIST so aufzeigen, daß sich auch der informationstechnische Laie ein Bild von den Möglichkeiten dieser Datenbank machen kann. Angepeilt werden insbesondere Personen, die an quantitativ-historischen Auswertungen interessiert sind.

* Informatikbulletin zur Informatik an höheren Mittelschulen des Kantons Bern, Nr. 16, Februar 1992. Bezugsquelle: Erziehungsdirektion des Kantons Bern, Amt für Unterricht, Sulgeneckstr. 70, CH-3005 Bern.

¹ Vgl. Kap. 2.2.2.

² Software Package for Social Sciences. Das 'X' bezeichnet eine bestimmte Programmversion.

³ Die vollständige Dokumentation kann gegen eine Schutzgebühr von Fr 20.- bezogen werden bei: Forschungsstelle für Regional- und Umweltgeschichte am Historischen Institut, 'Bernhist-Dokumentation', Unitobler, 3000 Bern 9.

⁴ Pfister, C., in Vorbereitung.

⁵ Imfeld, K., Häberli, P., in Vorbereitung.

1.1 Der Werdegang der Datenbank

Die Datenbank BERNHIST ist aus äußerst bescheidenen Anfängen heraus entstanden und ohne ein anfängliches Gesamtkonzept entsprechend den jeweils verfügbaren zusätzlichen Mitteln Schritt für Schritt ausgebaut worden. Am Anfang stand 1984 ein Pilotprojekt zur Erfassung von Taufen und Sterbefällen in den Kirchenbüchern. Die tabellarische Anordnung der Daten in den Originalquellen kam den technischen Gegebenheiten der damaligen EDV-Systeme entgegen und ermöglichte eine quellennahe Transkription. Dabei wurden nach dem Provenienzprinzip vorgegangen. Das heißt Einzelne Quellen wurden in ein EDV-System der BEDAG¹⁰ eingelesen¹¹ und mit dem Statistikpaket SPSSX ausgewertet. Daraus wurden neue Variablen berechnet, die in weiteren Dateien abgelegt wurden. Dieses Vorgehen bewährte sich, solange sich der Umfang des Datenbestandes auf eine übersichtliche Zahl von Quellen und Variablen beschränkte.

Je stärker jedoch der Datenbestand wuchs, desto schwieriger wurde es, die Übersicht zu behalten. Jede Datei enthielt mehrere Variablen, die entweder direkt aus einer Quelle übernommen wurden oder aber mit unterschiedlichen SPSSX-Programmen erzeugt worden waren. Auf diese Weise entstanden vielfältig verschachtelte Abhängigkeiten zwischen Dateien und Programmen. Dies machte es aufwendig, Daten aufzufinden, zu bearbeiten und wieder abzulegen. Mit der Zeit kannte sich bloß noch ein einziges Mitglied der Forschungsgruppe in dem fast undurchdringlichen Dateiendschangel aus.

In dieser Situation bot sich die Reorganisation der Datenbestände zu einer relationalen Datenbank als beste Lösung an. Den Anstoß dazu vermittelte das Berner Jubiläumsjahr 1991. Eine erste PC Version an einem speziellen Stand im Rahmen der jährlich stattfindenden Ausstellung BEA wurde zum Publikumserfolg. Lustbetont und neugierig ging das Messepublikum an die historisch-statistische Information heran, sei es, um Eckdaten der eigenen Wohngemeinde abzurufen, sei es, um erfahrene Facetten von »Vergangenheit« auf den Bildschirm zu holen oder sei es auch nur, um die Möglichkeiten der Software spielerisch zu erproben. Daraufhin wurde ein größeres Datenbankprojekt vom SEVA Lotteriefonds und von der Jubiläumstiftung der Schweizerischen Mobiliar Versicherungsgesellschaft unterstützt.

1.2 Anforderungsprofil der Datenbank

Die neue Datenbank wurde nach dem Pertinenzprinzip angelegt. Das heißt die einzelnen Quellen wurden in ihre Bestandteile, die Variablen, zerlegt und diese wurden nach thematischen Gesichtspunkten gegliedert und über die Zeit hin-

¹⁰ Bernische Datenverarbeitungs-AG.

¹¹ Die Quellen wurden am Computer mit einem Textsystem abgetippt und als Rohdatei gespeichert.

weg unter einheitlichen Begriffen zusammengefaßt, sofern dies angemessen war. Zu Recht bemängelt Peter Denley, daß bei dieser Vorgehensweise die Integrität, die interne Struktur und die individuelle Besonderheit der einzelnen Quelle ausgeblendet werden¹². Aber es war dies die einzige gängige Möglichkeit, das Computersystem selbst zum 'Organisator' der Datenbestände zu machen. Zur Neukonzipierung des gesamten Zugriffssystems wurde eine Datenbankumgebung evaluiert, die den folgenden Hauptkriterien entsprechen sollte:

- Der Zugriff sollte via Netzwerk von jedem IBM-kompatiblen PC aus mit standardisierter Software möglich sein.
- Mehrere Benutzer sollten gleichzeitig mit dem Datenmaterial arbeiten können.
- Die Datenbank sollte über eine einfach zu bedienende und gut verständliche Oberfläche abgefragt werden können.
- Die Kontrolle über den Benutzerkreis und Zuteilung von Zugriffsrechten an die einzelnen Benutzer sollte technisch einfach zu realisieren sein.
- Die Überwachung der Hardware, das Netzwerkmanagement, die Datensicherung und damit zusammenhängende Aufgaben sollten von den Informatikdiensten der Universität Bern erledigt werden können, da das Forschungsteam mit diesen Aufgaben nicht noch zusätzlich belastet werden konnte.

Diese Überlegungen führten zum Entschluß, die Daten vom BEDAG-Rechner auf die VAX der Informatikdienste der Universität Bern überzuführen. Dort standen alle wichtigen Entwicklungswerkzeuge zur Verfügung¹³, die ein solches Vorhaben überhaupt erst realisierbar machten. 1991 lief eine erste Version von BERNHIST auf dieser Anlage für Laien ohne spezielle EDV-Kenntnisse. Die vorliegende zweite Version wurde stark verbessert und so gestaltet, daß eine Abfrage auch über öffentlich zugängliche Terminals möglich ist. Dazu mußten folgende Möglichkeiten angeboten werden:

- Online-Verfügbarkeit des gesamten Datenmaterials.
- Gute Übersichtlichkeit mit Hilfe eines hierarchisch strukturierten Thesaurus.
- Einfache Zusammenstellung von Zeitreihen aus beliebigen Variablen.
- Räumliche Darstellung von Werteverteilungen (d.h. Import der Daten in ein Programm zur Herstellung von Karten).

¹² Denley, P., S. 38.

¹³ Für technisch Interessierte sei hier kurz auf die verwendeten Entwicklungswerkzeuge hingewiesen: BERNHIST läuft auf dem Betriebssystem VMS und basiert auf einer DEC-Rdb-Datenbank. Die Abfrage-Oberfläche wurde komplett in C geschrieben, die Datenbankoperationen erfolgen mit Embedded-SQL (Precompiler von DEC). Für die C-Programme standen verschiedene Bibliotheken (z.B. für die Programmierung der Menusteuerung) zur Verfügung, die - zum Teil als Public-Domain Routinen für MS-DOS entwickelt - von uns auf die VMS-Ebene transferiert wurden.

- Transfer von ausgewähltem Datenmaterial auf den eigenen PC¹⁴ oder auf benutzereigene Ressourcen im Grosssystem zur weiteren Verarbeitung.
- Import von extern erarbeiteten Daten, sei es im Rahmen von Lizentiatsarbeiten, Dissertationen oder Projektwochen an Mittelschulen.

Um dieses Pflichtenheft zu erfüllen, mußte ein streng standardisiertes, aber einfach verständliches Rohdatenformat festgelegt werden. Dieses ist in der Gesamtdokumentation¹⁵ von BERNHIST beschrieben. Dieses Format ist als Diskussionsgrundlage für ein universell anwendbares Vorgehen bei der Transkription von historisch- statistischen Quellen sowie der Anlage von statistisch bearbeitetem Material zu verstehen.

Ein gemeinsamer Datenpool für regionalgeschichtliche Forschungsarbeiten ist nur dann realisierbar, wenn alle Beteiligten ihre Ergebnisse in eine standardisierte Form bringen können, die eine Integration in die zugrunde liegende Datenbank rasch und einfach ermöglicht.

Die Möglichkeit zum Import von Daten ist eine wesentliche Voraussetzung für ein kostensparendes Recycling von forschungsproduzierten Dateien. Im Rahmen von Einzelarbeiten wird üblicherweise nur ein beschränkter Datenbestand ausgewertet. Nach dem Abschluß der Untersuchung verliert dieser für den Bearbeiter seinen Wert. Im besten Fall wird er publiziert und /oder eine Zeitlang archiviert; aber sofern er nicht in ein geeignetes Gefäß integriert werden kann, verkommt er früher oder später unvermeidlich zur Dateileiche.

Bei einem datenbankgestützten Vorgehen, wie es mit BERNHIST möglich wird, kann ein Großteil der benötigten statistischen Informationen digitalisiert aus der Datenbank übernommen werden, was die Formulierung entsprechend anspruchsvollerer Zielsetzungen erlaubt. Nach Abschluß der Untersuchung werden die Ergebnisse, soweit sie in quantifizierter Form vorliegen, in die Datenbank integriert und können von weiteren Bearbeiterinnen genutzt werden (vgl. *Abb. 2*)

Standards für die Wahrung der wissenschaftlichen Eigentumsrechte haben sich bei gedruckten Werken seit längerer Zeit durchgesetzt. Bei Datenbanken

¹⁴ BERNHIST ist so gestaltet, dass die auf einen IBM-kompatiblen PC transferierten Daten problemlos in Standardprogramme, die unter Windows laufen, importiert werden können. Besonders zu erwähnen ist Microsoft EXCEL Version 5, ein weitverbreitetes Programm zur tabellarischen und grafischen Präsentation sowie zur numerischen Verarbeitung von Daten. BERNHIST-Daten können mit diesem Programm einerseits darstellerisch aufbereitet werden - es stehen z.B. eine Vielzahl von Diagrammtypen zur Verfügung - oder auch in beschränktem Rahmen statistisch ausgewertet werden. EXCEL besticht jedoch nicht gerade durch einen grossen Umfang an statistischen Funktionen. Diesem Mangel kann durch Zusatzmodule abgeholfen werden. Z.B. DDU Software (Ed.), Astute Statistics Add-in for Microsoft EXCEL, The University of Leeds, Old Medical School, Leeds LS2 9JT, UK. Zum Transfer der Daten auf den lokalen PC vgl. Kapitel 3. Der Transfer von Daten auf einen Rechner des Typs 'Apple Macintosh' ist derzeit nicht möglich.

¹⁵ Vgl. Fussnote 7.

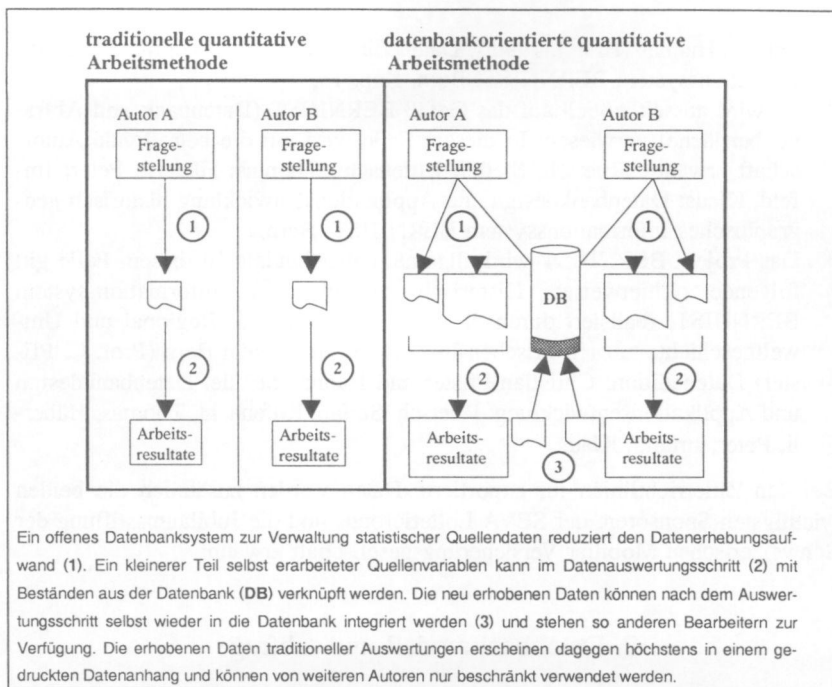


ABB. 2: Vorteile eines Datenbanksystems für die Erhebung und Verwaltung von statistischen Quelldaten. Quelle: Rüfenacht, T.; Salis Gross, C., 1993, S. 34.

besteht in dieser Hinsicht noch große Unsicherheit, und zwar deshalb, weil neben den inhaltlichen Aspekten auch die Konzeption und Ausprogrammierung der Datenbank selbst, also die Schaffung des Gefäßes, als eigenständige intellektuelle Leistung gewürdigt werden muß. Die für BERNHIST gültige Zitierrichtlinie unterscheidet zwischen den folgenden drei Fällen:

1. Es wird auf benutzte Datenbestände aus BERNHIST verwiesen, ohne das Datenbankgefäß speziell zu erwähnen: In diesem Fall wird die Autorschaft grundsätzlich gleich gehandhabt wie bei einem Sammelband, resp. einer Edition, wobei die Hauptkategorien des Thesaurus die »Kapitel« der Datenbank darstellen. Zitiert wird: Pfister, Christian; Schüle, Hannes (Hrsg.), 1990ff: Datenedition historisch geographisches Informationssystem BERNHIST. Bern. Alle auf Kosten des Projekts aufgenommenen und ausgewerteten Quellen werden unter dem Namen der Herausgeber zitiert. Dagegen haben Bestände, die im Rahmen von Einzelarbeiten quellenkritisch geprüft und ausgewertet wurden, den entsprechenden Verfasser zum Autor und werden unter dessen Namen in einer LIESMICH-Datei nach dem Transfer der Daten auf den lokalen PC zitiert. Beispiel: Thomas Rüfenacht, Curdin Salis Gross, Erwerbs Sektoren 1990, in: Pfister, Christian (Hrsg.),

Schule, Hannes (Hrsg.), 1990ff: Datenedition historisch-geographisches Informationssystem BERNHIST. Bern, Kap. 13.

2. Es wird ausschließlich auf das Gefäß BERNHIST (Datenbank und Abfrageoberfläche) verwiesen: In diesem Falle wird nur die betreffende Autorschaft erwähnt: Ruetsch, Stefan; Rüfenacht Thomas; Häberli, Peter, Imfeld, Klaus: Datenbankdesign und Applikationsentwicklung historisch geographisches Informationssystem BERNHIST. Bern.
3. Das Projekt BERNHIST wird allgemein abgehandelt: In diesem Falle gilt folgende Zitierweise: Historisch-geographisches Informationssystem BERNHIST, realisiert durch die Forschungsstelle für Regional und Umweltgeschichte am Historischen Institut der Universität Bern (Prof. C. Pfister) Datenedition: Christian Pfister und Hannes Schüle. Datenbankdesign und Applikationsentwicklung: Ruetsch, Stefan; Rüfenacht, Thomas; Häberli, Peter; Imfeld, Klaus.

Bei den Zitierrichtlinien für exportierte Daten werden zusätzlich die beiden wichtigsten Sponsoren, der SEVA Lotteriefonds und die Jubiläumsstiftung der Schweizerischen Mobiliar Versicherungsgesellschaft erwähnt

2. Datenbankmodell und -abfrage

2.1. Das Drei-Achsen-Modell

BERNHIST ist nicht auf spezifische Fragestellungen zugeschnitten, sondern ermöglicht einen universellen Zugang zu Daten, die auf der Basis der Einwohnergemeinden für das Territorium des Kantons Bern erhoben worden sind.

Die Datenbank basiert auf einem Drei-Achsen-Modell, das der dreidimensionalen Grundstruktur von Daten in historisch-statistischen Quellen entspricht. Jeder Beobachtungswert kann den folgenden drei Achsen zugeordnet werden¹⁶ (vgl. Abb. 3).

- Raum Geographischer Bezugsort.
- Thema Inhaltliche Aussage, die dem Wert zugeordnet ist.
- Zeit Zeitpunkte oder Zeitintervalle die mit dem Wert korrespondieren.

BERNHIST ist als relationale Datenbank¹⁷ konzipiert, die die computerinterne Organisation und Ablage der Daten selbständig regelt und eventuelle formale Konsistenzverletzungen registriert und meldet. Die Grundidee des relationalen Konzepts besteht für dieses Projekt darin, daß die physisch zusam-

¹⁶ Abbildung nach: Bahrenberg, G., Giese, E., Nipper, J., 1990⁷.

¹⁷ Eine sehr gute Abhandlung über verschiedene Datenmodelle unter besonderer Berücksichtigung des relationalen Konzepts bietet Vetter, M., 1986⁷. Mehr als Einführung zu empfehlen ist Date, C.J., 1981.

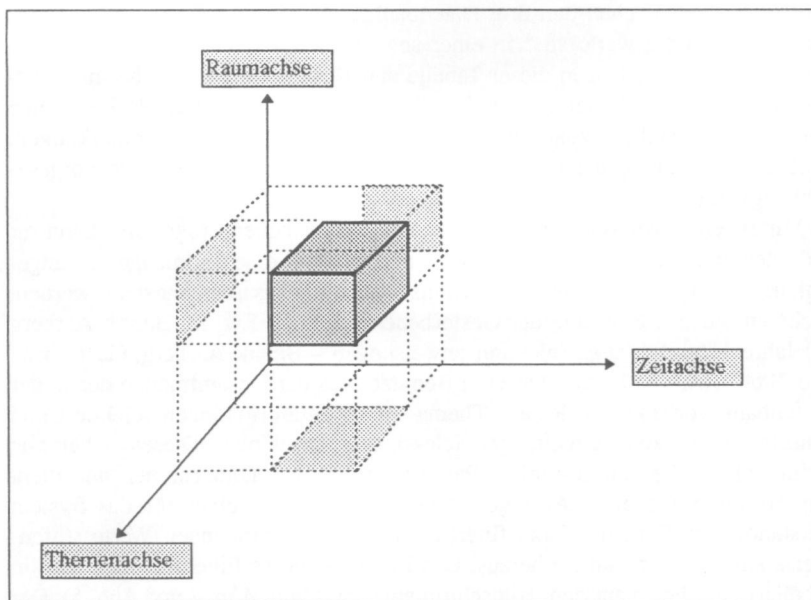


ABB. 3: Jeder Beobachtungswert in der Datenbank BERNHIST kann den drei Achsen Raum, Thema und Zeit zugeordnet werden.

menhängende, computerlesbare Information einer transkribierten Quelle in verschiedene, definitorisch klar voneinander abgegrenzte Teilbereiche zerlegt wird und diese Teilbereiche (Raum, Zeit und Thema) in je eine gesonderte **Datenbanktabelle** abgelegt werden. Diese Tabelle bildet eine **Entitätsmenge**, d.h. eine eindeutige Kollektion von Entitäten gleichen Typs¹¹, z.B. Namen von geographischen Räumen mit der dazugehörigen Raumkategorie. Jede **Entität** (Eintrag) in einer solchen Tabelle erhält einen **Schlüssel**¹², den das System automatisch vergibt, und auf den sich Informationen, die in anderen Tabellen lagern, wieder beziehen können. Die Daten können nur dann erfolgreich erfaßt und verwaltet werden, wenn sie konsistent sind. Dies verlangt, daß äquivalente Objekte immer mit den gleichen Begriffen beschrieben werden müssen. Es erfordert auch, daß in einem Thementhesaurus (siehe Kap. 2.2.2) keine Redundanzen (mehr als ein Beleg für einen einzelnen Begriff) vorkommen dürfen. Inkonsistenzen und Redundanzen verunmöglichen jede gültige Verwendung der Daten. Die Relationen zwischen den Datenbanktabellen werden durch ein Bezugssystem von verschiedenen Schlüsseln erzeugt und geregelt. Die drei Achsen der obigen

¹¹ Vetter, M., 1986, S. 37.

¹² Ein Schlüssel ist nichts anderes als eine Zahl, die vom Datenbankadministrator oder vom System selber der betreffenden Entität zugewiesen wird und diese dadurch eindeutig identifizierbar macht.

Abbildung entsprechen den drei Datenbanktabellen Raum, Thema und Zeit²⁰. Die Beobachtungswerte sind in einer sogenannten Wertetabelle (Observation Table) abgelegt. Jedem in dieser Tabelle enthaltenen Wert ist ein Raumschlüssel, ein Themenschlüssel, sowie eine Zeitinformaton, die in BERNHIST eine Jahreszahl darstellt²¹, zugeordnet. Weitere Tabellen sind für Anmerkungen, Quellenbezeichnungen und -nachweise sowie systeminterne Zwecke vorgesehen (vgl. Kap. 2.3).

Mittels eines von der Forschungsgruppe geschriebenen Programms kann für jede der drei Zuordnungskomponenten nach **Abb. 3** eine Teilmenge der angebotenen Parameter zu einer Selektionsanweisung zusammengestellt werden. Nehmen wir das Beispiele der Gestorbenen pro 1000 Ew. im Bezirk Aarberg im Jahre 1750. Die drei Selektionsanweisungen - Bezirk Aarberg, Gestorbene pro 1000 Ew., 1750 - werden vom Benutzer aus der Gesamtmenge der in der Datenbank vorhandenen Raum- Thema- und Zeitinformatonen selektiert und vom Programm zu einem einzigen Selektionsset kombiniert. Dieses enthält alle Schlüssel zu allen ausgewählten Parametern der drei Dimensionen und dient zur Formulierung einer Abfrage (Database-request) in einer für das System verständlichen Sprache. Dazu filtert es die übereinstimmenden Werte stufenweise aus der Wertetabelle heraus. Das Ergebnis dieses Filterprozesses wird in tabellarischer Form auf dem Bildschirm angezeigt (vgl. **Abb. 4** und **Abb. 5**). Der Programmteil zur Gewinnung einer Selektionsanweisung auf Benutzerebene wurde so gestaltet, daß dieser Vorgang möglichst intuitiv erfaßbar ist und keine schriftlichen Eingaben seitens der Benutzer erforderlich sind (siehe Kap. 2.2).

Der Benutzer kann die Datenbanktabellen zu Raum, Thema und Zeit in modifizierter Form auf den Bildschirm holen und durch einfaches Markieren der gewünschten Entitäten auf allen drei Dimensionen eine Selektionsanweisung zusammenstellen. Die folgende Abbildung versucht diesen Prozeß zu verdeutlichen. Es handelt sich dabei um nichts anderes als eine detaillierte, auf optische Zweidimensionalität reduzierte Variante von **Abb. 3**, wobei mit Hilfe von Pfeilen das Herausfiltern eines Wertes aus der Beobachtungstabelle verdeutlicht wird. Aus Gründen der Übersichtlichkeit wurde die Selektionsanweisung so gewählt, daß am Schluß nur ein einziger Wert aus der Wertetabelle herausgefiltert wird. In Wirklichkeit können mehrere tausend Werte gleichzeitig selektiert werden.

Nachdem die Datenbankabfrage die zu der Raum-Zeit-Thema Kombination gehörenden Werte aus der Beobachtungstabelle herausgefiltert hat, erzeugt

²⁰ Aus Gründen der Vereinfachung wird die zeitliche Dimension hier als Datenbanktabelle geführt, obwohl sie in Wirklichkeit nur ein Attribut der Wertetabelle (vgl. **Abb. 4**) ist.

²¹ Da die Zeitachse nur aus numerischen Werten besteht und ein Kontinuum darstellt, in dem jeder Wert eindeutig auf einen und nur einen Achsenpunkt projizierbar ist, können zeitliche Angaben gleichzeitig auch als Schlüssel dienen, d.h. sie müssen nicht wie die Raum- und Themainformation - mit einem numerischen Wert zur Identifikation versehen werden.

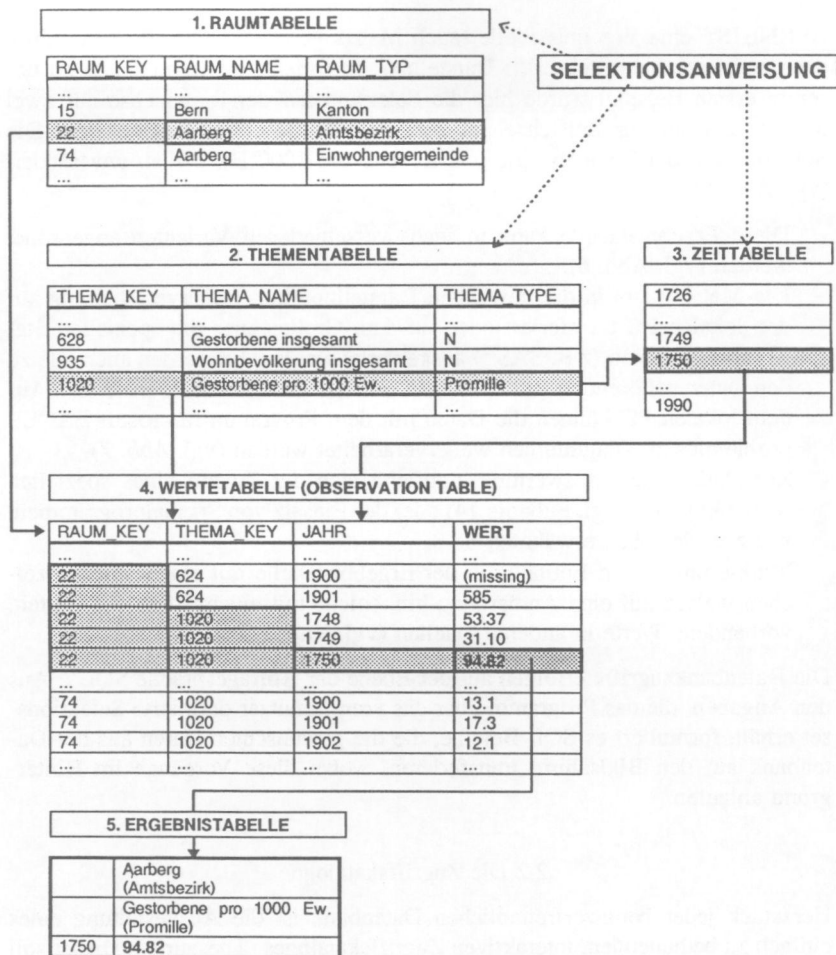


ABB. 4: Vereinfachtes Modell des Abfrageprozesses von BERNHIST anhand eines Beispiels. Aus dieser Abbildung werden die grundlegenden relationalen Zusammenhänge ersichtlich. Aus der Raum-, Themen- und Zeittabelle wird das Selektionsset gebildet, indem die jeweiligen Entitäten durch Druck auf die Leertaste ausgewählt und mit der Wertetabelle verknüpft werden. Die grauschattierten Bereiche aus der Wertetabelle werden vom Programm als Ergebnistabelle in Matrixdarstellung ausgegeben, wobei die Schlüssel für die Raum- und Themeninformation (RAUM_KEY, THEMA_KEY) wieder durch Klartext ersetzt werden. Auf der Abfrageoberfläche sind diese Schlüssel nicht sichtbar und hier nur zur Verdeutlichung des Prinzips abgebildet.

BERNHIST eine Ergebnistabelle (auch Matrix genannt) auf dem Bildschirm, die formal der untenstehenden Darstellung entspricht. Im Unterschied zu unserem ersten Beispiel wurde hier die Selektion auf der Raumachse auf zwei Räume und auf der Zeitachse auf zehn Jahre (1745-1755) ausgedehnt. Die Selektion auf der Themenachse (Gestorbene pro 1000 Ew.) blieb unverändert (vgl. *Abb. 5*).

- Diese Ergebnistabelle kann in sechs verschiedenen Varianten angeordnet werden (vgl. *Abb. 6*).
- Die Matrix kann in der gewählten Darstellung zur Weiterverarbeitung auf den lokalen PC transferiert oder auf dem Großrechner mit speziellen Statistikprogrammen (z.B. SAS²²) ausgewertet werden. Es werden auch alle zu den Daten gehörenden Anmerkungen und Quellenverweise transferiert. Auf dem lokalen PC können die Daten mit dem Programm Microsoft EXCEL problemlos zu Diagrammen weiterverarbeitet werden (vgl. *Abb. 7*).
- Zur statistischen Auswertung empfiehlt sich der Zukauf eines speziellen Statistikmoduls (vgl. Fußnote 14) oder der Einsatz von Spezialprogrammen, wie z.B. des oben erwähnten SAS.
- Markierungen von Entitäten in der Ergebnistabelle mit einem Ausrufezeichen weisen auf eine Anmerkung hin, solche mit einem Stern auf weitere vorhandene Werte in anderen Quellen (vgl. Kap. 2.3).

Die Datenbankzugriffe erfolgen auf der Ebene der Abfragesprache SQL²³. Aus den Angaben, die das Programm über das vom Benutzer definierte Selektionsset erhält, formuliert es SQL-Befehle, die die gewünschten Daten aus der Datenbank auf den Bildschirm transferieren, wobei diese Vorgänge im Hintergrund ablaufen.

2.2 Die Zugriffskataloge

Herzstück jeder benutzerfreundlichen Datenbank ist die Ausgestaltung eines einfach zu bedienenden, interaktiven Zugriffskataloges (Thesaurus). Dieser soll einen raschen Überblick über die vorhandenen Datenbestände ermöglichen. Dazu muß er hierarchisch strukturiert sein. Weiter soll er der interaktiven Formulierung einer Selektionsanweisung dienen. Der Benutzer soll sich seine Abfrage aus Informationen zusammenstellen können, die vorgefertigt auf dem Bildschirm erscheinen, so daß er nur noch die ihn interessierenden Bestandteile aus einer Liste auswählen muß. Er muß also selber keinen Text eingeben und

²² SAS, Statistical Analysis System ist ein Programmsystem zur Informationsverarbeitung und statistischen Datenanalyse. Eine gute Einführung bietet z.B.: Kähler, W.M., Schulte, W., 1987.

²³ Structured Query Language. Es handelt sich hier um eine häufig verwendete, auf verschiedenen Datenbanksystemen lauffähige Hochsprache, deren Syntax aber nur zum Teil standardisiert ist. Literatur: Marsch, J., Fritze, J., 1993.

y-Achse mit Zeitinformation	x1-Achse mit Rauminformation		x2-Achse mit Themeninformation	
	Aarberg (Amtsbezirk)	Burgdorf (Einwohnergemeinde)		
	Gestorbene pro 1000Ew. (Promille)			
1745	27.99	* 26.30		
1746	28.59	* 26.50		
1747	29.90	* 33.70		
1748	53.37	* 40.40		
1749	31.10	* 46.80		
1750	94.82	35.50		
1751	27.50	28.20		
1752	28.07	27.60		
1753	26.12	30.20		
1754	30.26	29.40		
1755	31.71	32.20		

ABB. 5: Ergebnistabelle in Form einer Matrix. Sie enthält die Werte der ausgewählten thematischen Variablen in den gewählten zeitlichen und räumlichen Dimensionen. Den Daten können wahlweise die quellennahen Terme oder die vereinheitlichten Begriffe zugeordnet werden (vgl. Kap. 2.2.2). Zu den Ausrufezeichen und Sternchen vgl. Kap. 2.3.

Variante	y-Achse	x ₁ -Achse	x ₂ -Achse
(A)	Thema	Raum	Zeit
(B)	Thema	Zeit	Raum
(C)	Raum	Thema	Zeit
(D)	Raum	Zeit	Thema
(E)	Zeit	Raum	Thema
(F)	Zeit	Thema	Raum

ABB. 6: Die sechs möglichen Anordnungsformen der BERNHIST-Matrix. In Abb. 5 wurde die Darstellungsvariante (E) gewählt.

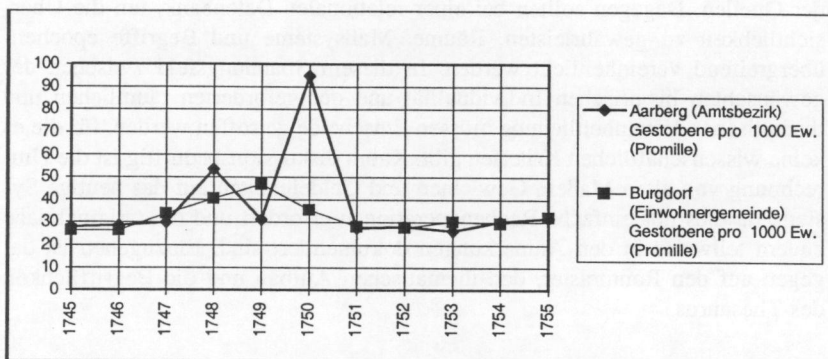


ABB. 7: Grafische Umsetzung von Abb. 5 mit dem Programm Microsoft EXCEL. Die extreme Sterblichkeit im Bezirk Aarberg ist durch die größte Bevölkerungskatastrophe Berns seit den Pestjahren, die Rote Ruhr von 1750 bedingt (Pfister, 1989).

auch keine Systemkenntnisse auf der Ebene der SQL-Datenbanksprache mitbringen.

Über die Thesauren wird die gewünschte Komplexitätsreduktion vorgenommen. Dies bedingte einen hohen Programmieraufwand. Zur Formulierung der Selektionsanweisungen für alle drei Dimensionen (vgl. *Abb. 3* und *Abb. 4*) verfügt BERNHIST über zwei Zugriffskataloge:

- **Raumkatalog:** Er stellt die administrative Gliederung des Kantons in hierarchischer Form dar.
- **Themenkatalog:** Er faßt zusammengehörige Variablen in Kategorien zusammen. Diese können wiederum einer Oberkategorie zugeordnet werden. Zwischen der unstrukturierten Thementabelle in der Datenbank (vgl. *Abb. 4*) und dem Thementhesaurus ist ein Vereinheitlichungsschritt enthalten, von dem in Kapitel 2.2.2 die Rede sein wird.

In *Abb. 4* wurde schon einmal verdeutlicht, wie die Datenauswahl in BERNHIST zu geschehen hat. Über die dort eingeführten Begriffe Raumentabelle, Thementabelle und Zeittabelle wird jetzt noch das System der Zugriffskataloge gelegt, welches diese Tabellen mittels für den Benutzer unsichtbaren Zusatzinformationen nach bestimmten strukturellen Gesichtspunkten auf einer Metaebene ordnet und auf dem Bildschirm darstellt.

Mit der Schaffung von BERNHIST werden historische Daten mit Strukturen und Prozeduren der Informatik in Beziehung gesetzt. Dabei stellt sich das Problem, wie historische Inhalte datenbankgerecht und benutzerfreundlich aufbereitet werden können, ohne daß dabei ihr Wesensgehalt allzusehr verfremdet wird. Die Geschichtsschreibung ist darauf angelegt, dem individuellen Charakter jeder Epoche gerecht zu werden. Dazu gehören als Rahmenbedingungen u.a. die territoriale Ordnung, Münzen und Maße, Sprache und Begrifflichkeit der Quellen. Dagegen sollten bei einer relationalen Datenbank, um die Übersichtlichkeit zu gewährleisten, Räume, Maßsysteme und Begriffe epochenübergreifend vereinheitlicht werden. In diesem Spannungsfeld zwischen der gewünschten historischen Individualität und der geforderten räumlichen und thematischen Vereinheitlichung müssen Entscheide getroffen werden, für die es keine wissenschaftlichen Kriterien gibt. Kaum diskussionsbedürftig ist die Umrechnung von alten Maßen, Gewichten und Geldeinheiten auf das heutige System, da dies nur einfache Rechenoperationen erfordert und die Originalmaße zudem teilweise in den Anmerkungen dokumentiert sind. Einzugehen ist dagegen auf den Raumraster, den thematischen Aufbau und die Begrifflichkeit des Thesaurus.

2.2.1 Raumkatalog und Raumstandardisierung

Der Raumkatalog von BERNHIST enthält eine hierarchisch gegliederte Liste aller Gemeinden, Kirchgemeinden, Amtsbezirke und Landesteile des Kantons Bern. Die hierarchische Gliederung wird wie folgt erzeugt:

- Jede einzelne Raumeinheit (Raum-Entität) besteht aus einer Raumbezeichnung, einem Raumcode und einer Raumkategorie (vgl. *Abb. 9*). Durch die Raumkategorie ist die vertikale Zuordnung der Raum-Entität in der Raumhierarchie eindeutig festgelegt. Raumkategorien (in absteigender hierarchischer Ordnung) können sein: Kanton, Landesteil, Amtsbezirk, Kirchgemeinde, Einwohnergemeinde. In dieser Reihenfolge wurde die Hierarchie definiert (vgl. *Abb. 8*).
- Raum-Entitäten mit gleichartigen Raumkategorien bilden eine *Raumebene*.
- Von jeder Raum-Entität ist bekannt, welcher übergeordneten Raumebene und welcher übergeordneten Raum-Entität sie angehört. So kann die Position innerhalb der Raumhierarchie bestimmt werden.

Diese Aussagen führen nur dann zu einer eindeutig reproduzierbaren Zuordnung, wenn der Raumraster über den gesamten Zeitraum hinweg stabil bleibt, über den in der Datenbank Informationen gespeichert sind. Im Kanton Bern hat es aber insbesondere im 19. Jahrhundert außerordentlich viele Gebietsveränderungen und -Verschiebungen gegeben. Weil die Implementation einer zeitabhängigen Raumstruktur²⁴ zu aufwendig gewesen wäre, wurde während der Projektierungsphase von BERNHIST beschlossen, die Raumstruktur des Kantons Bern aus dem Jahre 1980 zum verbindlichen Standard zu erklären. Eine Ausnahme bilden einzig die Kirchgemeinden, für die aus Gründen der besseren Vergleichbarkeit mit den Verhältnissen des 18. und des 19. Jahrhunderts der Zeitpunkt 1900 als Standardisierungshorizont gewählt wurde. Die Standardisierung bedingt, daß Daten aus transkribierten Quellen zuerst durch ein SAS-Programm laufen müssen, welches die Daten auf den Raumstandard von 1980 bringt, bevor sie in die Datenbank integriert werden können. Berücksichtigt werden dabei Eingemeindungen, Wechsel von Gemeinden zu anderen Amtsbezirken sowie Namenswechsel von Gemeinden. Reine Gebietsverschiebungen zwischen Raum-Entitäten, die durch die Festlegung neuer Grenzlinien entstanden sind, konnten nicht berücksichtigt werden. In diesen Fällen ist das Auftreten von quantitativen Sprüngen in den Datenreihen möglich. Diese Sprünge werden aber durch Anmerkungen zu der betroffenen Raum-Entität kommentiert.

Die so definierte Raumstruktur kann als baumartiger Graph abgebildet werden, der sich in eine Anzahl von Unterbäumen verzweigt, die sich ihrerseits weiter verzweigen. Eine solche Baumstruktur ist programmtechnisch einfach zu implementieren.

In der vorliegenden Version von BERNHIST muß die Raumhierarchie in Tabellenform dargestellt werden. Im Raumkatalog werden die verschiedenen Raumebenen durch das Maß des horizontalen Einzugs der jeweiligen Raumentität dargestellt (vgl. *Abb. 9*).

²⁴ Über dieses wichtige Problem wird in einer späteren Publikation noch ausführlich die Rede sein. Es ist an unserer Forschungsstelle ein Ansatz entwickelt worden, der dieses Problem vollständig löst. Der Aufwand, diesen zu konkretisieren und implementieren hat sich aber als zu teuer erwiesen.

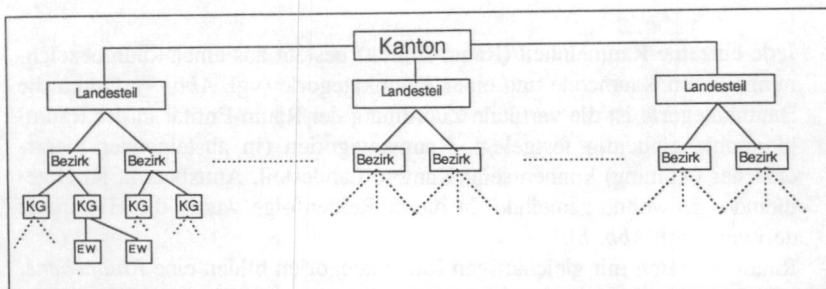


ABB. 8: Die hierarchische Raumstruktur, dargestellt als Baum. Legende: KG=Kirchgemeinde, EW=Einwohnergemeinde.

Raumcode	Raumbezeichnung	Raumkategorie
99	Bern	Kanton
...
52	Seeland	Landesteil
1	Aarberg	Amtsbezirk
101	Aarberg	Kirchgemeinde
5301	Aarberg	Einwohnergemeinde
...
5306	Lyss	Einwohnergemeinde (!)
...
4	Biel / Bienne	Amtsbezirk
...
51	Südiura	Landesteil

ABB. 9: Ausschnitt aus dem Raumkatalog der erweiterten Datenbank als Beispiel für die konkrete Darstellung auf dem Bildschirm. In der öffentlichen Version von BERNHIST sind Kirchgemeinden und Landesteile nicht enthalten. Eine Raum-Entität (Zeile) ist definiert durch die drei Komponenten Raumcode, Raumbezeichnung und Raumkategorie. Die Raumcodes der Einwohnergemeinden entsprechen dem Code des Bundesamtes für Statistik zu denen aus logistischen Gründen die Zahl 5000 addiert wurde. Das Ausrufezeichen hinter der Raum-Entität "Lyss - Einwohnergemeinde" weist auf eine Anmerkung zu diesem Raum hin (vgl. Kap. 2.3).

ABB. 9: Ausschnitt aus dem Raumkatalog der erweiterten Datenbank als Beispiel für die konkrete Darstellung auf dem Bildschirm. In der öffentlichen Version von BERNHIST sind Kirchgemeinden und Landesteile nicht enthalten. Eine Raum-Entität (Zeile) ist definiert durch die drei Komponenten Raumcode, Raumbezeichnung und Raumkategorie. Die Raumcodes der Einwohnergemeinden entsprechen dem Code des Bundesamtes für Statistik zu denen aus logistischen Gründen die Zahl 5000 addiert wurde. Das Ausrufezeichen hinter der Raum-Entität "Lyss - Einwohnergemeinde" weist auf eine Anmerkung zu diesem Raum hin (vgl. Kap. 2.3).

Abb. 9 stellt einen bestimmten Ausschnitt der BERNHIST-Abfrageoberfläche auf dem Bildschirm dar. Mit Hilfe der Positionsbalken (hier grauschilder) dargestellt) kann der Benutzer die gesamte Raumhierarchie durchkämmen und sich durch einfachen Tastendruck eine Auswahl der gewünschten Raum-Entitäten zusammenstellen. Um die Suche zu erleichtern, verfügt der Katalog über verschiedene Zusatzfunktionen:

- Suchfunktion nach bestimmten Raumbezeichnungen oder -codes.
- Aus- und einblenden ganzer Raumbereichen (z.B. aller Kirchgemeinden)
- Auswahl ganzer Unterbäume (z.B. aller Einwohnergemeinden eines Amtsbezirks).

- Suchfunktion nach bestimmten Raumbezeichnungen oder -codes.
- Aus- und einblenden ganzer Raumebenen (z.B. aller Kirchgemeinden)
- Auswahl ganzer Unterbäume (z.B. aller Einwohnergemeinden eines Amtsbezirks).

Abschließend noch zwei Bemerkungen:

In der öffentlichen Version von BERNHIST sind die Raumebenen 'Landesteil' und 'Kirchgemeinde' nicht enthalten, um das Auswahlverfahren zu vereinfachen. Letztere wurden als statistische Basiselemente des Staates 1833 von den Einwohnergemeinden abgelöst und sind in ihrem räumlichen Umfang heute nur noch wenigen gegenwärtig. Fast alle Daten aus dem 18. und 19. Jahrhundert sind aber auf die Ebene der Einwohnergemeinden umgerechnet worden (Stand 1980). Die letzte Bemerkung gilt den Raumcodes: Da die Datenbank zwei voneinander verschiedene Gemeindesysteme beinhaltet, wurde zum BFS-Code²⁵ der Einwohnergemeinden 5000 addiert, um Überschneidungen mit den dreistelligen Codes der Kirchgemeinden²⁶ zu vermeiden.

2.2.2 Themenkatalog und Themenstandardisierung

Der Themenkatalog (Thementhesaurus) enthält eine in Kategorien gegliederte, begrifflich vereinheitlichte Form der Thementabelle aus der Datenbank. In der Thementabelle der Datenbank sind die aus der transkribierten Quelle stammenden quellennahen Variablen, die sogenannten *Quellenterme*, völlig unstrukturiert abgelegt²⁷. Diese können nicht in den Thementhesaurus übernommen werden, sondern müssen vorgängig vereinheitlicht werden und zwar aus folgenden Gründen:

- Zur Wahrung der Konsistenz müssen äquivalente Objekte über größere Zeiträume hinweg immer mit denselben Begriffen beschrieben werden.
- Nur bei begrifflicher Konsistenz lassen sich nach dem in *Abb. 11* gezeigten Muster auf einfache Weise Zeitreihen zusammenstellen.
- Manche Quellenbegriffe sind nur für Spezialisten verständlich.

Wesentlich schwieriger ist es, die erforderliche sachliche Vergleichbarkeit zu gewährleisten. Je weiter wir in der Zeit zurückschreiten, desto stärker verändert sich die verwendete Terminologie und die Beziehung vom 'Wort' zur 'Sache'. Da tauchen Quellenbegriffe auf, die ohne kritischen Kommentar nicht verständlich sind, und manche historische Erscheinungen kennen keine Entsprechung mehr in der Gegenwart. Ehe Quellenbegriffe entsprechend dem heutigen Sprachverständnis »übersetzt«, d.h. einem vereinheitlichten Term zugeordnet

²⁵ Dieser Code ist ersichtlich aus: Bundesamt für Statistik Schweiz (Hg.) 1986«'6.

²⁶ Die Kirchgemeinden wurden wie folgt codiert: Die ersten zwei Ziffern entsprechen dem Code des Amtsbezirks (von 1-30 durchnummeriert, inkl. Bezirke auf dem Gebiet des heutigen Kantons Jura), die letzten zwei bilden eine Laufnummer, deren Wert sich aus der alphabetischen Reihenfolge der Raumnamen ergibt.

²⁷ Eine absolut quellentreue Übernahme der Begriffe aus Quellen in tabellarischer Form ist in vielen Fällen nicht möglich, da die Begriffe zum Teil aus verschiedenen in der Quelle enthaltenen Spaltenbezeichnungen zusammengesetzt werden müssen. Der Begriff »Quellenterm« bezeichnet demnach nur eine möglichst quellennahe Begrifflichkeit.

werden dürfen, hat sich die methodische Kritik mit der Entstehung und dem Inhalt der Quelle auseinanderzusetzen und die Vergleichbarkeit abzuklären. So wurde, um ein Beispiel herauszugreifen, bis in die Mitte des 19. Jahrhunderts einzig der Akt der Taufe, nicht aber die Geburt, amtlich verzeichnet. Der Eintrag in den Taufrodel diente zugleich der zivilen und der kirchlichen Registrierung. Da ein Teil der Kinder vorher starb, war die Zahl der Taufen stets geringer als jene der Geburten. Weder vom institutionellen Kontext noch von der Sache her ist damit die terminologische Gleichsetzung von »Taufen« und »Geburten« angebracht. Die Taufen sind nach einem Schätzverfahren auf Geburten hochgerechnet worden, um sie mit diesen vergleichbar zu machen. Das rechnerische Vorgehen beruht auf gewissen Annahmen und enthält damit Elemente der Interpretation. Doch erscheint die Zusammenfassung der beiden Termini unter dem gängigen Begriff »Geburten« vom Gesichtspunkt der Benutzerin her gerechtfertigt. Würden diese beiden Begriffe nämlich getrennt im Thesaurus geführt, müßten Taufen und Geburten zur Erstellung langer Zeitreihen vorgängig zusammengefaßt werden. Dies würde Zeit und Rechenaufwand kosten und den Benutzer zwingen, sich mit dem hinter der Operation stehenden inhaltlich-terminologischen Problemen auseinanderzusetzen. Da klar anerkannte Regeln in diesem Bereich fehlen, ist die Zusammenfassung von Quellbegriffen zu vereinheitlichten Begriffen problematisch und in manchen Fällen diskussionswürdig. Hier bedarf es noch einer erheblichen Anstrengung der Geschichtswissenschaft.

Die Transparenz ist dadurch gewährleistet, daß der hinter dem vereinheitlichten Term stehende Quellenbegriff sichtbar gemacht werden kann. So wird, um bei unserem Beispiel zu bleiben, in der Periode vor 1876 »Taufen«, zwischen 1876 und 1929 »Geburten«, nach 1930 »Lebendgeburten« als Quellenbegriff angezeigt. Aus Kostengründen ist dieses Nebeneinander von Quellenbegriffen und »Standardterms« nicht durchgängig umgesetzt werden. Eine flexiblere Zugangsweise könnte bei einem späteren Ausbauschritt durch die automatische Zuweisung von beliebigen Synonymbegriffen zu den vereinheitlichten Begriffen im Themenkatalog programmiert werden, wie dies in kommerzieller Software möglich ist. Auf der Ebene der Variablen sind die Quellennachweise vollständig, so daß in dieser Hinsicht die Forderung nach wissenschaftlicher Transparenz erfüllt ist.

Andererseits verlangt das wissenschaftliche Gebot der Transparenz, daß die maschinenlesbaren Daten anhand der Quelle überprüft werden können. Dies ist nur möglich, wenn auch die quellennahen Bezeichnungen in der Datenbank gehalten werden. Dieses Dilemma zwischen Quellennähe und Konsistenz wurde wie folgt angegangen:

1. Es wird eine Liste von vereinheitlichten Begriffen geschaffen, die sich so nahe wie möglich an jene des Bundesamtes für Statistik anlehnen (diese waren jedoch selber im Lauf der Zeit erheblichen Veränderungen unterworfen).

2. Die Begriffe **werden nach** Begriffsgruppen in Kategorien zusammengefaßt. Die Kategorien haben wiederum Begriffscharakter, so daß daraus übergeordnete Kategorien gebildet werden können. Es entsteht also - wie in Kapitel 2.2.1 beschrieben - ein hierarchisches System, das von der Forschungsgruppe selber angelegt worden ist (vgl. *Abb. 10* und *Abb. 11*).
3. Durch einen einheitlichen Aufbau der Begrifflichkeit wurde erreicht, daß eine logische Gliederung der Begriffe innerhalb der Kategorien bei der elektronischen Sortierung erhalten bleibt. Da sich Fachleute selten auf eine einheitliche Verwendung von Begriffen einigen können, ist die vorliegende Lösung als Diskussionsvorschlag zu betrachten.
4. Jedem **vereinheitlichten** Begriff können eine Anzahl von Quellentermen zugeordnet werden, die jederzeit abrufbar sind.
5. Nach der Datenbankabfrage muß die Darstellungsmatrix (vgl. *Abb. 5*) die Themeninformation sowohl in quellenaher wie auch in vereinheitlichter Form darstellen können.

Die meisten Daten liegen nur für die Jahre vor, in denen die entsprechenden Zählungen stattfanden. Um die Zähljahre schon bei der Auswahl sichtbar zu machen, und damit Leersuchen oder eine Ergebnistabelle mit zahlreichen fehlenden Werten zu vermeiden, wurde die zeitliche Dimension in den thematischen Zugriffskatalog integriert, so daß überhaupt nur Thema-Zeit Kombinationen gewählt werden können. In einer ersten Stufe des Auswahlverfahrens werden alle vorhandenen Kategorien angezeigt von denen dann wahlweise eine aufgeklappt wird (hier grau markiert, vgl. *Abb. 10*):

01	Bevölkerung	Kategorie
08	Erwerbsleben	Kategorie
03	Gebäude und Wohnungen	Kategorie
09	Land-, Alp- und Forstwirtschaft	Kategorie
07	Politik	Kategorie
02	Raum und Umwelt	Kategorie
13	Versicherungen, Soziale Gliederung und Fürsorge	Kategorie
06	Unternehmen und Arbeitsstätten	Kategorie

ABB. 10: Die Hauptkategorien des Thementhesaurus in der öffentlichen Version von BERNHIST. Die Kategorien sind jene des Bundesamtes für Statistik, die teilweise etwas erweitert werden mußten.

Der Inhalt der aufgeklappten Kategorie wird wiederum in Tabellenform auf dem Bildschirm angezeigt. Hier taucht der vereinheitlichte Begriff »Wohnbevölkerung insgesamt« auf, von dem zu bestimmten Jahren Zahlen verlangt werden sollen. Das Programm zeigt alle in Frage kommenden Jahre an, zu denen Daten vorliegen. Wahlweise kann die Kontrolle auch noch auf die Raumachse ausgedehnt werden, so daß nur Jahre angezeigt werden, zu denen für das betreffende Thema innerhalb der vom Benutzer definierten räumlichen Se-

lektionsanweisung Daten existieren. Die Jahreszahlen werden wiederum in einer separaten Tabelle dargestellt und können ausgewählt werden (starke Graufärbung, vgl. *Abb. 11*):

0111	Eheschliessungen	Kategorie
0101	Geburten	Kategorie
0102	Gestorbene	Kategorie
0109	Haushalte	Kategorie
0110	Wanderung	Kategorie
	Wohnbevölkerung insgesamt	N
0103	Wohnbevölkerung nach Alter und Geschlecht	Kategorie
0105	Wohnbevölkerung nach Heimatklasse	Kategorie
0106	Wohnbevölkerung nach Konfession (Religion)	Kategorie
0107	Wohnbevölkerung nach Muttersprache	Kategorie
0104	Wohnbevölkerung nach Zivilstand	Kategorie

JAH(R) WÄHLEN
...
1850
1856
1860
1870
1880
...

ABB. 11: Die Unterkategorien der Hauptkategorie "Bevölkerung" im Thementhesaurus von BERNHIST. Auf dieser Ebene ist außerdem die Variable "Wohnbevölkerung" angesiedelt. Die Darstellung veranschaulicht die Thema-Zeit-Koppelung in BERNHIST: Es werden nur jene Jahre angezeigt, für die in der Datenbank tatsächlich auch Werte enthalten sind.

Auf diese Weise können auch mehrere Thema-Zeit-Kombinationen gleichzeitig selektiert werden. Das Programm kombiniert sie danach mit der vom Benutzer festgelegten Rauminformation (Raumkriterium) und fragt die Wertetabelle der Datenbank nach korrespondierenden Werten ab. *Abb. 11* zeigt übrigens auch, daß Kategorien über mehrere Ebenen verschachtelt sein können.

2.3 Anmerkungsapparat und Quellenverweise

Die aktuelle Version von BERNHIST wurde mit umfangreichen Möglichkeiten zur Integration eines Anmerkungsapparates ausgestattet. Den wichtigsten Datenbanktabellen wurden separate Anmerkungstabellen zugeordnet, deren Inhalt über die Benutzeroberfläche abgefragt werden kann. Der Anmerkungsapparat wurde programmtechnisch mit dem Raum- und Themenkatalog sowie mit der Ergebnistabelle verbunden. Das heißt Wenn eine Anmerkung vorhanden ist, wird die betreffende Entität mit einem Ausrufezeichen markiert. Durch Tastendruck läßt sich die Anmerkung in einem separaten Bildschirmfenster darstellen.

Wenn BERNHIST-Daten auf den lokalen PC oder auf das eigene Konto auf dem Großrechner exportiert werden, werden alle Anmerkungen, die mit diesem Datensatz in Zusammenhang stehen, in einer separaten Textdatei zwingend mitexportiert. Der Benutzer kann demnach nur während der Formulierung der Selektionsanweisungen selber entscheiden, ob er die Anmerkungen sehen will oder nicht. Bei einem allfälligen Export der Daten dagegen werden immer alle damit verbundenen Anmerkungen mitgeliefert.

Das Anmerkungskonzept von BERNHIST basiert wiederum auf dem Prinzip der Kategorisierung. Es werden verschiedene Kategorien von Anmerkungen geschaffen, von denen jede einer oder mehreren Datenbanktabellen zugeordnet

ist (vgl. *Abb. 12*). Damit kann der Gültigkeitsbereich einer Anmerkung sukzessive eingeschränkt werden. Beispiele: Anmerkungen zum Raumkatalog (vgl. *Abb. 9*), welche auf **Gebietsveränderungen** hinweisen, werden unabhängig von den ausgewählten Themen immer dargestellt. Thema-Zeit-Anmerkungen werden jedoch nur dann sichtbar, wenn das betreffende Thema in Verbindung mit dem betreffenden Jahr gewählt worden ist (vgl. *Abb. 11*). Unser Beispiel in *Abb. 7* zeigt eine typische Möglichkeit zur Anwendung von Thema-Zeit-Anmerkungen: Das Thema »Gestorbene pro 1000 Ew.«, verbunden mit dem Jahr 1750 kann mit der Anmerkung versehen werden, daß 1750 die Mortalitätswerte im ganzen Kantonsgebiet generell höher waren, da in diesem Jahr die Rote Ruhr wütete. Wenn noch punktuellere Ereignisse vorliegen, etwa ein schlecht lesbarer und deshalb interpretationsbedürftiger Wert in einer handschriftlichen Quelle, können Anmerkungen zu einzelnen Werten definiert werden. In der Ergebnistabelle werden diese Werte mit einem Ausrufezeichen versehen (vgl. *Abb. 5*).

Anmerkungen können nicht interaktiv eingegeben werden, sondern sind Bestandteil einer RDF-Datei (siehe Kapitel 3), d.h. sie werden bei der Transkription einer Quelle oder bei der Integration statistischer Auswertungen definiert und eingegeben. Es ist jedoch auch möglich, nachträglich eine RDF-Datei zu erstellen, die Anmerkungen zu bereits in der Datenbank vorhandenen Begriffen und Daten enthält. *Abb. 12* faßt alle in BERNHIST vorhandenen Anmerkungskategorien zusammen, die in einer RDF-Datei definiert sind und auf der Programmoberfläche interaktiv abgerufen werden können.

Räumlich-thematisch verknüpfte Anmerkungen sind allerdings nicht möglich, d.h. es gibt keinen Anmerkungstyp, der sich auf eine Relation zwischen einem bestimmten Raum und einem bestimmten Thema bezieht.

Anmerkungen können nicht nur typisiert sondern auch nach inhaltlichen Kriterien klassifiziert werden. Attribute können sein:

- Qualität
- Erklärung
- Quelle
- Methode

Dieses Klassifikationssystem bietet die Möglichkeit, bestimmte Anmerkungsklassen ein- oder auszublenden. Die BERNHIST-Benutzeroberfläche bietet zu diesem Zweck ein Menu an, in welchem jede der vier Klassen aktiviert bzw. deaktiviert werden kann. Es werden nur Anmerkungen angezeigt, deren Klasse aktiviert ist Beispiel: Wer sich nur für erklärende Anmerkungen interessiert, kann die restlichen drei Klassen deaktivieren, so daß in allen fünf Anmerkungskategorien nur erklärende Anmerkungen angezeigt werden.

In der Ergebnistabelle (vgl. *Abb. 5*) liefert BERNHIST neben den Anmerkungen auch Quellenverweise. Diese können zu jedem Wert auf dem Bildschirm dargestellt werden, indem der Cursor²³ auf den entsprechenden Wert

²³ Der Begriff **Cursor** bezeichnet eine Markierung auf dem Bildschirm, die die aktuelle

Anmerkungskategorie	Erklärung
Raum-Anmerkung:	Zu einzelnen Raum-Entitäten (vgl. Abb. 9) können Anmerkungen definiert werden, was in der Tabellendarstellung durch Ausrufezeichen hinter der Raumkategorie angezeigt wird. Zum Beispiel Lyss: 1867 ging das Gebiet der Schulgemeinde Werdthof, das administrativ zu Lyss gehörte, zu Kappelen, 1957 das Gebiet Worbengarage von Kappelen zu Lyss.
Thema-Anmerkung:	Zu jedem vereinheitlichten Begriff im Themenkatalog (vgl. Abb. 10) können Anmerkungen definiert werden. Beispiel: Dem Thema „Zugtiere: Anzahl insgesamt“ wurde eine Anmerkung zugeordnet, welche Tierarten unter den Begriff „Zugtiere“ fallen, nämlich „Pferde, Ochsen und Kühe“.
Thema-Zeit-Anmerkung:	Jede Thema-Zeit-Kombination im Themenkatalog kann mit Anmerkungen versehen werden. Beispiel: 1750 wütete die Rote Ruhr im Kanton Bern, deshalb waren die Mortalitätswerte in diesem Jahr stark erhöht. Dem Thema „Gestorbene pro 1000 Ew.“ verbunden mit dem Jahr 1750 wird demnach eine Anmerkung zugeordnet, die diesen Peak in der Mortalitätskurve (vgl. Abb. 7) erklärt.
Quellenterm-Anmerkung²⁸:	Zu jedem quellennahen Begriff können - analog zur Thema-Anmerkung oben - Anmerkungen vergeben werden. Im Themenkatalog (vgl. Kap. 2.2.2) können zu jedem vereinheitlichten Begriff die Quellenterme in einem separaten Bildschirmfenster dargestellt werden. Ein Ausrufezeichen hinter dem Quellenterm bezeichnet dabei das Vorhandensein einer Anmerkung. Beispiel: Der vereinheitlichte Begriff „Pferde (Munche)“ im Themenkatalog von BERNHIST wurde dem Quellenterm „Munche“ zugeordnet. Eine Anmerkung zu diesem Quellenterm erklärt, was „Munche“ bedeutet, nämlich „kastrierte Hengste“.
Wert-Anmerkung:	Diese Anmerkungskategorie hat den feinsten Definitionsraster, d.h. es können Anmerkungen zu jedem in der Wertetabelle (vgl. Abb. 4) enthaltenen Wert definiert werden. Die Werte werden in der Ergebnistabelle (vgl. Abb. 5) mit einem Ausrufezeichen versehen, falls eine Anmerkung definiert wurde. Dieser Anmerkungstyp ist z.B. dann sinnvoll, wenn ein räumlich und thematisch isolierter Sprung in einer Datenreihe der Interpretation bedarf oder wenn ein Wert in einer handschriftlichen Quelle schlecht lesbar war, so dass verschiedene Deutungen möglich wären.

ABB. 12: Die Anmerkungskategorien von BERNHIST. Die Anmerkungen sind den Zugriffskatalogen oder der Ergebnistabelle zugeordnet. Ihr Vorhandensein wird mit einem Ausrufezeichen markiert. Die Anmerkungen können in separaten Bildschirmfenstern dargestellt werden und werden beim Export der Daten in eine separate Anmerkungsdatei geschrieben.

positioniert wird und der Quellenverweis via Menu abgerufen wird. Beispiel: Zum Wert für die Raum-Thema-Zeit-Kombination Aarberg (Amtsbezirk) - Gestorbene pro 1000 Ew. - 1750 erscheint der Quellenverweis »Berechnungen BERNHIST«, d.h. es handelt sich hier um einen Wert, der von einem Statistikprogramm errechnet wurde. Es ist auch möglich, daß zu einer Raum-Thema-Zeit- Kombination unterschiedliche Werte aus mehreren Quellen existieren. Da die

Ergebnistabelle wegen ihrer Zweidimensionalität jedoch nur einen Wert pro Raum-Thema-Zeit- Kombination darstellen kann, wird in einem solchen Fall der Wert aus der ersten gefundenen Quelle angezeigt Dieser Eintrag wird in diesem Fall mit einem Stern markiert (vgl. *Abb. 5*). Wenn hier der Quellenverweis abgerufen wird, werden alle gefundenen Quellen aller gefundenen Werte dargestellt. Beispiel: Das Thema Wohnbevölkerung 1988 im Amtsbezirk Signau besitzt folgende Quellenverweise:

- Wert 24393, in: Statistische Datenbank der Schweiz STATINF (BFS)
- Wert 24436, in: Berechnungen BERNHIST.

Beim Export der Daten auf den PC werden alle Quellenverweise und Mehrfachwerte in der ebenfalls mitexportierten Anmerkungsdatei abgespeichert.

3. Von der Quelle zur Datenbank

Zum Schluß soll noch kurz darauf eingegangen werden, wie eine Quelle Ihren Weg in die Datenbank findet und wie Forschungsergebnisse reintegriert werden können (vgl. *Abb. 2*). Der wichtigste Teil dieses Prozesses bildet ein spezielles Dateiformat, das Rohdatenformat (im folgenden RDF genannt). Es handelt sich dabei um einen Text mit formal standardisiertem Aufbau und streng definierten Syntaxregeln, der daraufhin angelegt ist, die verschiedenen historisch-statistischen Quellen möglichst vollständig erfassen und mit eigenen Ergänzungen (Anmerkungen) versehen zu können. Jede historische Quelle muß in ein solches RDF umgewandelt werden, ehe sie in die Datenbank integriert werden kann, wobei meist zwei oder noch mehr Schritte notwendig sind (vgl. *Abb. 13*). Danach kann das entstandene Rohdatenfile mit einem speziellen Programm in die Datenbank eingelesen werden. Die Schaffung eines leicht beschreibbaren Standardformates ist notwendig, weil so einerseits auch externe Forschungsergebnisse ohne aufwendige Umwandlungsarbeit sofort in der Datenbank abgelegt werden können und andererseits ein hoher Grad an Automatisierung und Übersicht erreicht wird. Wenn allgemein verständliche Regeln aufgestellt werden können, wie Dokumente aufgebaut sein müssen, um Eingang in unsere Datenbank zu finden, können auch Personen mit minimalen EDV-Kenntnissen

Bearbeitungsposition darstellt. Mit den Pfeiltasten auf der PC-Tastatur kann diese Markierung in jede beliebige Richtung verschoben werden.

dieselben erstellen. Ein solches Regelwerk befindet sich in der detaillierten Dokumentation zu BERNHIST²⁹. Um ein Chaos zu vermeiden, darf neues Material nur an zentraler Stelle von einem mit dem Programm vertrauten Datenbankadministrator eingelesen werden. *Abb. 13* soll den Weg der Quelle in die Datenbank grob verdeutlichen.

4. Erstellen von Kartogrammen mit dem Programm MapViewTM

4.1 Einleitung

Die kartographische Darstellung ist ein geeigneter Weg, um die große Datenmenge, die BERNHIST anbietet (vor allem, wenn alle 400 Gemeinden gezeigt werden sollen), übersichtlich und verständlich zu präsentieren. Die Forschungsstelle für Regional- und Umweltgeschichte der Universität Bern hat deshalb ein Instrumentarium entwickelt, um die BERNHIST-Daten in Karten umsetzen zu können.

Für die Kartenherstellung am PC haben wir uns für das Programm MapViewer³⁰ entschieden. Dieses Produkt läuft unter Microsoft Windows³¹ und ist deshalb kompatibel zu vielen verbreiteten Text-, Tabellenkalkulations- und Grafikprogrammen. Nach der relativ aufwendigen Arbeit des Erstellens einer Grundkarte können mit relativ geringem Aufwand BERNHIST-Daten importiert und in ansprechende Farbkarten umgesetzt werden. Das Resultat ist immer am Bildschirm sichtbar und kann sehr einfach editiert und gestaltet werden. Die erzeugten Karten können dann direkt gedruckt, oder in Grafik-, Textverarbeitungs- bzw. Desktop publishing-Programmen weiterverwendet werden.

4.2 Die Grundkarte (Base Map)

Die Base Map ist eine Karte der Gemeinde- und Amtsgrenzen des Kantons Bern auf dem Stand von 1980 (aber ohne Laufental). Sie besteht aus verschiedenen Schichten («Layers»), die jeweils separat bearbeitet, übereinandergelegt und ein- oder ausgeblendet werden können. Neben Layers für Seen, Punkte und Ortsnamen sind vier verschiedene Layers mit Amtso- oder Gemeindegrenzen erstellt worden. Demnach können Karten mit vier verschiedenen Flächenmustern angefertigt werden:

- Amtsbezirke
- Amtsbezirke: Fläche unter 1400 M.ü.M. (~~im~~ **Oberland**)

²⁹ Vgl. Fussnote 9.

³⁰ Map Viewer ist ein eingetragenes Warenzeichen der Golden Software Inc., 809 14th Street, Golden, CO 80401- 1866, USA.

³¹ Microsoft Windows ist ein eingetragenes Warenzeichen der Microsoft Corporation.

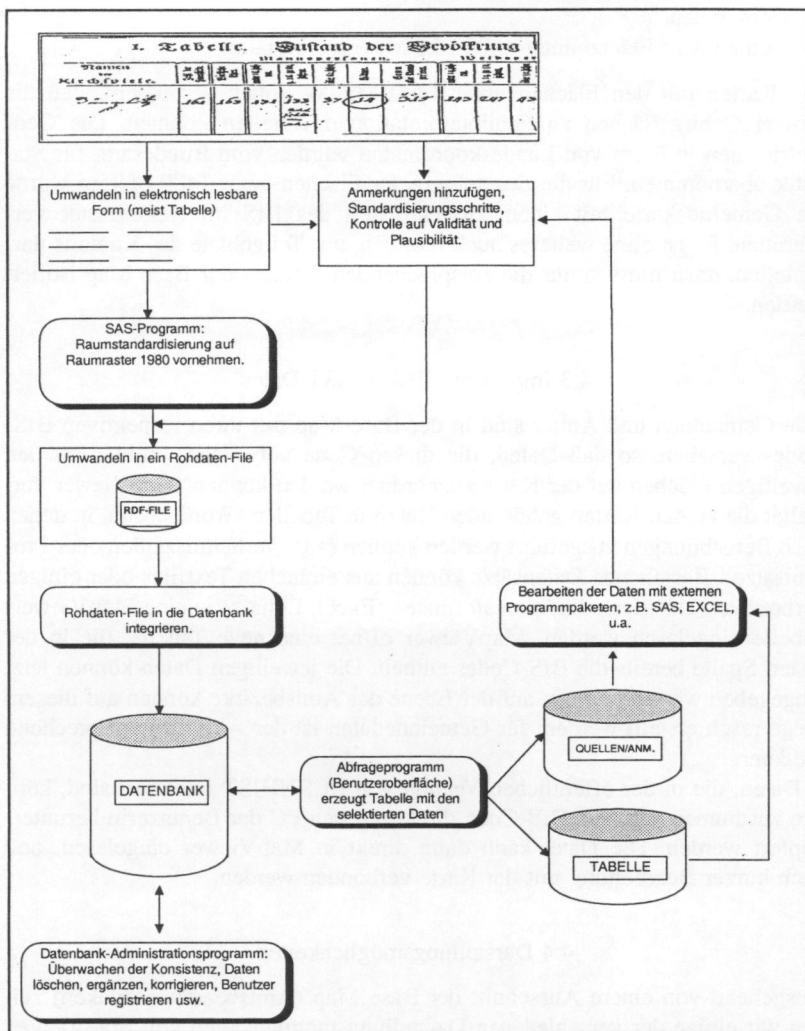


ABB. 13: Der Weg der Quelle in die Datenbank BERNHIST. Die Abbildung visualisiert das Gesamtsystem BERNHIST mit seinen peripheren Bestandteilen. Die abgerundeten Kästchen stellen - wo nichts anderes vermerkt ist - Programme dar, die von der Forschungsgruppe geschrieben wurden. Bei der handschriftlichen Quelle handelt es sich um die Ergebnisse der Volkszählung 1764 für die Kirchgemeinde Bümpliz.

- Gemeinden
- Gemeinden: Räche unter 1400 M.ü.M. (im Oberland)

Die Karten mit den Flächen unter 1400 M.ü.M. sollen verhindern, daß die großen Gebirgsflächen zu Fehlinterpretationen verleiten könnten. Die Geometriedaten in Form von Landeskoordinaten wurden vom Bundesamt für Statistik übernommen. Für die Darstellung der Flächen unter 1400 Metern wurde die Gemeindekarte mit einem Höhenmodell des BfS im Hektarraster verschnitten. Es ist ohne weiteres auch möglich, nur Teilgebiete des Kantons darzustellen, dazu müssen nur die entsprechenden Flächen der Base Map isoliert werden.

4.3 Import der BERNHIST-Daten

Alle Gemeinden und Ämter sind in der Base Map mit ihren respektiven BFS-Codes versehen, so daß Daten, die diesen Code aufweisen, problemlos den jeweiligen Flächen auf der Karte zugeordnet werden können. Map Viewer verwaltet die zu den Karten gehörenden Daten in Tabellen (Worksheets), in denen auch Berechnungen ausgeführt werden können (z.B. Verhältniszahlen oder Prozentsätze). Bestehende Datensätze können aus einfachen Textfiles oder einigen verbreiteten Tabellenkalkulationsformaten (Excel, Lotus³²) in eine Map Viewer-Tabelle eingelesen werden. MapViewer öffnet eine neue Tabelle, die in der ersten Spalte bereits die BFS-Codes enthält. Die jeweiligen Daten können jetzt eingegeben werden. Karten auf der Ebene der Amtsbezirke können auf diesem Wege rasch erstellt werden, für Gemeindedaten ist der Aufwand entsprechend größerer.

Daten, die in der öffentlichen Version von BERNHIST integriert sind, können zusammen mit dem BFS-Code direkt auf den PC der Benutzerin herunterkopiert werden. Die Datei kann dann direkt in MapViewer eingelesen, und nach kurzer Bearbeitung mit der Karte verbunden werden.

4.4 Darstellungsmöglichkeiten

Ausgehend von einem Ausschnitt der Base Map (Amtsbezirk Interlaken) zeigen wir einige der verschiedenen Darstellungsmöglichkeiten von MapViewer. Für viele Fragestellungen am besten geeignet ist die Flächenkarte (Hatch Map). Farben und Füllmuster können frei gewählt werden. Die Einteilung der Klassengrenzen kann automatisch (gleiches Intervall oder gleiche Anzahl pro Klasse) oder individuell geschehen.

Zur Darstellung mehrerer Variablen eignet sich die Balken-Karte (Bar Map). Die Balken können generell in der Mitte der Flächen oder individuell positioniert werden. Breite, minimale und maximale Höhe und Farbe der Balken kön-

³² Lotus ist ein eingetragenes Warenzeichen der Lotus Development Corporation.

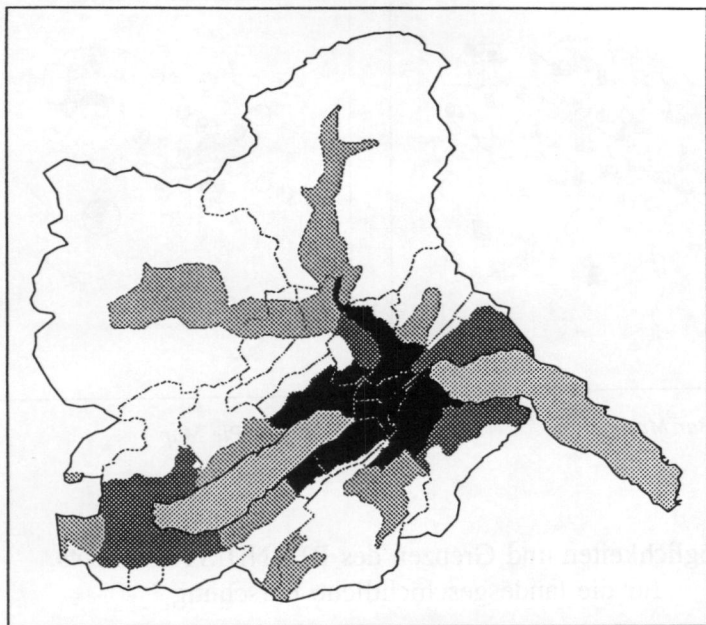


ABB. 14: Hatch Map mit Grundkarte Gemeindeflächen unter 1400 Meter.

ABB. 14: Hatch Map mit Grundkarte Gemeindeflächen unter 1400 Meter.

nen individuell gewählt werden. Wenn mehrere Variablen zusammen 100 % ergeben (z.B. Erwerbstätige nach Sektoren), bietet sich das Kuchendiagramm (Pie Map) als geeignete Darstellungsform an. Die Größe des »Kuchens« ist proportional zur Größe des Summenwertes der Variablen. Maximale und minimale Größe der Kuchen sowie deren Farbe können bestimmt werden. Die Positionierung der Kuchen erfolgt automatisch in der Flächenmitte oder kann frei gewählt werden.

4.5 Legenden

Für jeden Kartentyp (außer Base Map und Prism Map") kann MapViewer automatisch eine Legende erzeugen. Diese bleibt immer aktuell, paßt sich also Änderungen in der Karte oder an den Daten an. Für die definitive Kartenversion kann ihre Gestaltung den individuellen Wünschen angepaßt werden (vorher ist es nicht empfehlenswert, weil dann Änderungen in der Karte nicht mehr automatisch auf die Legende übertragen werden).

³³ Prism Map: Ohne Abbildung.

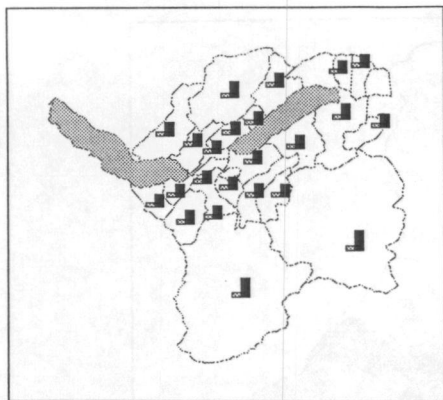


ABB. 15: *Bar Map*

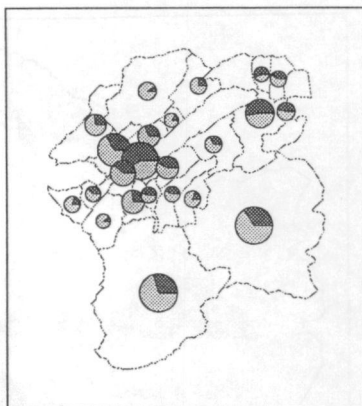


ABB. 16: *Pie Map*

5. Möglichkeiten und Grenzen des BERNHIST-Konzepts für die landesgeschichtliche Forschung

Alle Theorien über längerfristigen Wandel argumentieren quantifizierend, sei es im Bereiche der materiellen Kultur, sei es im Bereiche der Denk- und Verhaltensweisen. Stets postulieren sie ein Mehr oder ein Weniger, handle es sich um Geburten, um Ernährung, um religiöse Bindungen oder um Disziplin, um nur einige Beispiele zu erwähnen. Solche Veränderungen sind nicht nur als Aussagen über vergangene Wirklichkeiten, sondern auch als Rahmenbedingungen des Handelns von Individuen und Gruppen wissenswert. Eine Datenbank zur Darstellung raum-zeitlicher Daten und zur Erforschung von quantitativen Zusammenhängen im Sinne der historischen Sozialforschung¹⁴ kann dabei auf vierfache Weise zur Bereicherung, Präzisierung und Objektivierung von historischen Aussagen beitragen:

1. Durch Erschließung eines bedeutenden Korpus' von historisch-statistischen Quellen. Die vorwiegend aus Zahlen bestehenden Schriftquellen sind bisher erst punktuell in die Geschichtsschreibung einbezogen worden, weil ihre Aufbereitung mit großem Aufwand verbunden ist und ihre Interpretation außerhalb der Tradition des Faches hegt. Die Erschließung dieses Materials bietet Anstoß zur Bearbeitung neuer, bisher unzugänglicher oder wenig attraktiver Fragestellungen¹⁵. Zur Erschließung gehört die Prüfung der inneren Konsistenz des Materials: Der Historiker ist sich gewöhnt, Quellen

¹⁴ Ruloff, D., 1985.

¹⁵ Botz, G., S. 56.

aller Art mit einem gesunden Mißtrauen zu begegnen und diese vor Verwendung auf ihre Tendenz, Genauigkeit und Repräsentativität kritisch zu hinterfragen. Quantitativen Datenmaterial, vor allem dann, wenn es von staatlicher Seite zusammengetragen wurde, wird dagegen oft ein blindes Vertrauen entgegengebracht. Sei es, daß vergessen wird, daß auch diesem Material gegenüber die üblichen quellenkritischen Vorbehalte am Platz sind¹⁶. Sei es, daß der Wille oder die Möglichkeit zur Prüfung der inneren (quantitativen) Konsistenz solcher Daten fehlen.

2. Statistische Daten bieten die Möglichkeit, impressionistische Äußerungen erzählender Quellen gleichsam von einem Außenstandpunkt aus zu überprüfen, der von individueller Wahrnehmung unabhängig ist. Was immer die Geschichtsschreibung für wahr halten möchte, sie muß sich die Frage gefallen lassen, ob es vereinbar sei mit dem, was die Statistik an Deutungen gestattet. Aus einer sorgfältigen gegenseitigen Überprüfung von Aussagen quantitativer und narrativer Quellen werden letztlich Ergebnisse gewonnen, die wissenschaftlich konsistenter sind und längerfristig gültig bleiben als solche, die nur auf einer einzigen Quellengattung beruhen¹⁷.
3. Zur graphischen und kartographischen Umsetzung können ausgewählte Daten mit geringem Aufwand aus BERNHIST exportiert, auf andere Systeme transferiert und dort mit kommerzieller Software in ihrer räumlichen und zeitlichen Dimension dargestellt werden. Dies ermöglicht eine grobe Prüfung der Konsistenz, und oft vermitteln die Darstellungen Ansätze zur Formulierung oder Prüfung von Hypothesen. Wichtig ist die Feststellung, daß durch die räumliche Darstellung ein Weg eröffnet wird, um raumbezogene Theorien (z. B. »Zentrum - Peripherie«) oder Konzepte (z.B. »Ökotypus«) zu operationalisieren und in die zeitspezifischen Fragestellungen der Geschichtswissenschaft zu integrieren¹⁸.
4. Bei der theoriegesteuerten Verknüpfung von Material aus verschiedenen Quellen zu Metaquellen: Die ältere historisch sozialwissenschaftliche Forschung beschränkte sich auf die Auswertung einzelner historisch-statistischer Quellen, wobei die möglichen Fragestellungen zwangsläufig durch den Inhalt der Quelle vorgegeben waren. Dagegen ist in BERNHIST eine Vielzahl von Quellen gespeichert, die Schlüsseldaten für mehrere Teilsysteme der Gesellschaft - Bevölkerung, Wirtschaft, Umwelt, Politik - enthalten, wobei die einzelnen Variablen so aufbereitet sind, daß sie außerhalb des Kontexts ihrer Quelle angesprochen, ausgewählt und beliebig mit anderen vernetzt werden können. Damit eröffnet sich die Möglichkeit zur Schaffung von Metaquellen. Peter Becker versteht darunter Konstrukte, die Informationen aus verschiedenen isolierten Quellenbeständen vereinigen, aber keine direkte quellenmäßige Entsprechung haben¹⁹. Als Beispiel einer

¹⁶ Sharlin, A., zit. in Ruloff, D., 1985, S. 90.

¹⁷ H. Siegentaler in Ritzmann-Blickenstorfer, 1995, in Vorbereitung.

¹⁸ Mücke, H., 1988.

¹⁹ Becker, P., 1987.

komplexen Metaquellen kann die Schätzung der gesamten Nahrungsmittelproduktion (unter Einschluß von Getreide, Hackfrüchten, Fleisch und Fisch, Milch und Milchprodukten, Obst und Nüssen, Gemüse, Wein und Sammel Früchten) für den Kanton Bern im Zeitraum 1850 erwähnt werden⁴⁰.

Bei aller Euphorie darf die historiographische Bedeutung der Datenbank nicht überschätzt werden. Limitierend dürfte bei Fragestellungen einmal die räumliche Begrenzung auf den Kanton Bern wirken; immerhin verbindet sich mit der »Publikation« von BERNHIST die Hoffnung, daß trotz der Knappheit der Mittel in absehbarer Zeit kompatible Datenbanken für weitere Kantone geschaffen werden. Grenzen setzt im weiteren die Beschränkung des Gefäßes auf serielle, quantitative und tabellarische Quellen. Dadurch bleibt seine Reichweite auf einen Kreis von Fragestellungen beschränkt, die sich mit diesen Quellentypen ausleuchten lassen.

6. Auswahlbibliographie

BERNHIST verdankt seine Entstehung in einer embryonalen Phase einer Initiative der bernischen Erziehungsdirektion (Abteilung Unterricht). Später ist die Datenbank mit Unterstützung des Schweizerischen Nationalfonds ausgebaut worden. Das heutige Konzept ist mit Mitteln des SEVA- Lotteriefonds und der Jubiläumsstiftung der Schweizerischen Mobiliar Versicherungsgesellschaft realisiert worden. Wolfgang Polasek, Universität Basel; Ulrich Pfister, Universität Zürich; und Hannes Schüle, Universität Bern, haben zu diesem Artikel wertvolle Anregungen beigegeben.

Bahrenberg, G.; Giese, E.; Nipper, J., 1990: Statistische Methoden in der Geographie. Bd. 1. Univariate und bivariate Statistik. Stuttgart.

Becker, P., 1987: Formen und Möglichkeiten der Standardisierung bei Metaquellen. in: Hausmann, F. e. al. (Hg.), Datennetze für die historischen Wissenschaften? Probleme und Möglichkeiten bei Standardisierung und Transfer maschinenlesbarer Daten. Graz 1987, S. 18-27.

Botz, G., 1984: Was gewinnt die Geschichtswissenschaft durch die Quantifizierung. In: Nagl- Docekal H., Wimer, F. (Hg.), Neue Ansätze in der Geschichtswissenschaft, Wien, S. 48- 70.

Bundesamt für Statistik Schweiz (Hg.) 1986: Amtliches Gemeindeverzeichnis der Schweiz, in: Amtliche Statistik der Schweiz, Nr. 049, Bern.

Date, C.J., 1981: An Introduction to Database Systems, Addison Wesley Publishing Company, o.O.

Denley, P., 1994: Models, Sources and Users: Historical Database Design in the 1990's. In: History and Computing 6/1, S. 33-43.

⁴⁰ Pfister, C, Schüle, H., 1988.

- Imfeld, K.; Häberli, P., in Vorbereitung: Dokumentation zur Datenbank BERNHIST. Seminararbeit am Historischen Institut der Universität Bern.
- Kahler, W.M.; Schulte, W., 1987: SAS für Anfänger. Einführung in das Programmsystem, Braunschweig, Wiesbaden.
- Marsch, J.; Fritze, J., 1993: SQL: Eine praxisorientierte Einführung. Braunschweig, Wiesbaden.
- Mücke, H., 1988: Historische Geographie als lebensweltliche Umweltanalyse. Studien zum Grenzbereich zwischen Geographie und Geschichtswissenschaft. Frankfurt.
- Pfister, C., 1989: Der rote Tod im Kanton Bern: Demographische Auswirkungen und sozio- hygienisches Umfeld von Ruhrepidemien im 18. und 19. Jahrhundert unter dem Einfluß einer umweltorientierten Medizin, in: Saladin, P.; Schauffelberger, H.J; Schläppi, P. (Hgg.), 1989: 'Medizin' für die Medizin. Arzt und Ärztin zwischen Wissenschaft und Praxis. Festschrift für Hanes G. Pauli, Basel und Frankfurt am Main.
- Pfister, C., 1995: Regionen im Strom der Modernisierung. Bevölkerung, Wirtschaft und Umwelt im Kanton Bern 1700-1914. Bern.
- Pfister, C., in Vorbereitung: Bevölkerung, Wirtschaft und Umwelt im Kanton Bern, 1700-1990, Archiv des Historischen Vereins des Kantons Bern.
- Pfister, C.; Schüle, H., 1988: BERNHIST. Eine Quellen- und Datenbank wird zum Ausgangspunkt einer neuen Landesgeschichte für den Kanton Bern. In: Kaufhold, K. H., Schneider, J. (Hgg.), Geschichtswissenschaft und elektronische Datenverarbeitung, Wiesbaden.
- Ritzmann-Blickenstorfer, H. (Hg.), 1995: Historische Statistik der Schweiz, Zürich, in Vorbereitung.
- Ruloff, D., 1985: Historische Sozialforschung, Stuttgart.
- Ruetsch, S.; Rüfenacht T.; Häberli, P.; Imfeld, K.: Datenbankdesign und Applikationsentwicklung historisch geographisches Informationssystem BERNHIST. Bern.
- Rüfenacht, T.; Salis Gross C., 1993: Der Eisenbahnbau und die räumliche Verteilung der Wirtschaft im Kanton Bern 1850-1910, Lizentiatsarbeit in Schweizer Geschichte, Bern.
- Vetter, M., 1986: Aufbau betrieblicher Informationssysteme mittels konzeptioneller Datenmodellierung, Stuttgart
- Vetter, M., 1988: Strategie der Anwendungssoftwareentwicklung. Planung, Prinzipien, Konzepte, Stuttgart